Arnaud Sahuguet, Ph.D.

New York, NY 10027

+1 646-246-0942

arnaud.sahuguet@gmail.com

linkedin.com/in/sahuguet

twitter.com/@sahuguet

medium.com/@sahuguet

sahuguet.github.io

SUMMARY ↑

My areas of expertise include data – big, small, open, geographic – management and analysis, architecture of complex systems, product design and research & development.

My passion is to invent, architect and build products that leverage technology to solve meaningful problems and have a large social impact. I aim at empowering people and organizations to be more productive and collaborative through innovation.

I have developed some strong working relationships with the academic research community, the non-profit sector and innovators at the city and government levels. I have also acquired some domain expertise for agriculture and climate data and data-driven applications and models.

Education | Experience | Add. information | Publications | External engagements | Patents

EDUCATION ↑

University of Pennsylvania, Philadelphia, PA   1996–2001
Ph.D. in computer science (advisor: Val Tannen).

Ecole Normale Supérieure, Univ. Paris VI, Paris, France   1995–1996
M.S. in algorithms and cryptography.

Ecole Nationale des Ponts et Chaussées, Paris, France   1994–1996
M.S. in civil engineering.

Ecole Polytechnique, Palaiseau, France   1991–1994
B.S. in engineering with major in computer science.

EXPERIENCE ↑

TORBI LLC   3-2016–present
Owner, New York,
Through TORBI LLC, I provide various consulting services including product coaching, product design, software development, hiring, due diligence, etc.

Gro Intelligence   2/2016–3/2024
Consultant, Interim SVP Eng, Chief Product Officer, SVP Product, New York,
As Chief Product Officer, I oversaw data ingestion, data representation, data modeling, data discovery, data serving and the various apps we created. I was leading Gro's ontology team. I was heavily involved in data engineering. And I was in charge for Gro's AI strategy for both internal and customer-facing applications.

EXOR Seeds   4/2019–3/2021

Venture Partner, New York,

- sourcing.
- technical due diligence.

Jacobs Technion-Cornell Institute   3/2018–8/2020

Project lead for the Urban Tech Hub, New York,

- leading the development of the Urban Tech hub where we will train entrepreneurs and technologists to understand the structure of cities and the needs of communities they are building solutions for.
- conducted an extended market research with more than 120 interviews (academia, public and private sectors) to identify opportunities in this space.
- helped the Cornell development team secure a $15M gift to launch the hub.
- coordinated the efforts of a faculty task force to design the structure and curriculum of a 2 year MS program.
- prepared the paperwork for the approval of a new concentration with New York State Education Department (approved in 9/2019).
- hired the hub's first executive director.
- co-organized the New Day at the MTA conference (September 2019).

The Foundry @ Cornell Tech   9/2015–12/2018

Director, New York,

- helping faculty and students move from ideas to products and companies.
- recruited and managed a team of 6 full-time staff members.

The Governance Lab @ NYU   6/2014–8/2015

Chief Technology Officer and Principal, New York,

- responsible for all IT and technology for a lab of 20 people.
- coordinating the technology aspect of all lab's projects and proposals.
- recruited and managing the engineering team (2 people).

Google Inc.   3/2006–3/2014

Product manager, Google.org, New York, 5/2011–4/2014

- co-founded the Big Civic Data team (5 people) and the Colaboratory Data Science project (launched as Google Colab in October 2017) within Google Research; I conducted interviews with various cities to understand real needs; I identified the iPython notebook as a starting point for an integrated, interactive, visual and collaborative data science platform; I advocated the idea to internal teams (Chrome, Docs, Cloud) as well as externally; the project has been open sourced at part of project Jupyter.
- co-founded OneToday, a mobile platform for micro-donations, launched in May 2013; highest rated Google app on PlayStore; I was involved in all facets of the launch, including legal, marketing, partnership; I personally enrolled more than 500 non-profits to create enough project inventory for the launch.
- product lead for our effort on civic services, open data and e-government solutions; I co-authored a schema.org vocabulary for government services working with the Whitehouse Office of Science and Technology Policy (OSTP) and New York City.
- led our internal efforts on child protection as the Google liaison with the National Center for Missing and Exploited Children (NCMEC).

Product manager, Geo team, New York, 5/2008–5/2011
- co-architect for the next-generation unified infrastructure for Google Maps; I coordinated 20 engineers across 4 different teams.
- designed & implemented analytical and bug tracking tools for geo data support teams.
- 3 related patents filed.

Product manager, Google Research and speech, New York, 08/2006–11/2008
- defined strategy and product roadmap for speech efforts including dialogue systems, speech transcription and speech synthesis (foundations for Google Voice Search).
- launched Google Audio Indexing (GAudi) for the 2008 US political campaign.
- launched GOOG-411, Google first voice product (US and Canada); secured funding from CEO and coordinated marketing campaign (billboard, taxi, etc.); negotiated deal with phone manufacturer for "GOOG-411 phone" (CES 2008).

20% projects and others, New York, 03/2006–3/2014
- internal champion for Google barcode efforts; co-founded the ZXing open source library for Android, used by BarcodeScanner, a featured T-Mobile application that appeared on TV ads; 2 related patents.
- launched "nearby street intersection" feature for non-visual maps applications.
- Google *trusted interviewer*, with more than 400 interviews conducted.

Bell Labs Research, Lucent Technologies　　　　　　　　　　　　　2001–2006
Member of technical staff, Murray Hill, NJ, 09/2001–03/2006
- Lucent representative for 3GPP on Generic User Profile; key contributor for TS 22.240, TS 23.240 and TS 29.240.
- researcher focusing on data federation, personalization, identity management and converged services; publications and tutorials presented at various database and web conferences.

Tropea Inc.　　　　　　　　　　　　　　　　　　　　　　　　　1999-2002
Founder, CTO, Philadelphia, PA, 1999-2002
- Founded company to commercialize W4F web extraction tools.
- Sold licenses to end-user customers and integrator companies.
- Released technology as open source project after joining Bell Labs.

DigiCash bv　　　　　　　　　　　　　　　　　　　　　　　　　1996
Intern, Amsterdam, NL, 03/1996–07/1996
- Implemented an electronic voting prototype for TeleDemocratie project, with Berry Shoenmakers, David Chaum and Marcel Bullinga.

Parisbas Capital Markets　　　　　　　　　　　　　　　　　　　1994
Intern, London, UK, 09/1994–12/1994
- Rewrote and implemented of various pricing tools, with Eric Auld and Bruno Dupire.

French Armed Forces　　　　　　　　　　　　　　　　　　　　　1991-1992
Sous-lieutenant at 11e R.C., Carpiagne, France, 03/1992–08/1992
Aspirant at EABC, Saumur, France, 10/1991–02/1992

ADDITIONAL INFORMATION ↑

- US and French citizen.

- Spoken languages: French (native), English (fluent), Spanish (good).

- Programming languages: Python, Javascript, SQL; functional languages.

- Sports: soccer, marathon (3h20 PR), ultra-marathon (Black Rock City, 2011), triathlon.

SELECTED PUBLICATIONS ↑

h-index=19; total citations=1,935; full list of publications available from DBLP.

- Circular Data for a Circular City: Value Propositions for Mobility, The Circular City Research Journal, Volume I, Spring 2019.

- Smarter New York City - How City Agencies Innovate, Columbia University Press, 2018. (chapter 11 on Vision Zero).

- Les Big Data à découvert, CNRS Editions, 2017. (part V, chapter 10 with Amélie Marian).

- Everything* You Always Wanted to Know About ICOs But Were Afraid to ask, Medium, Oct 2017.

- Data, Responsibly, Dagstuhl Seminar 16291, July 2016.

- A "calculus" for Open Data, with David Sangokoya, Medium, Feb 2015.

- Open Civic Data: Of the People, By the People, For the People, with John Krauss, Luis Palacios and David Sangokoya, IEEE Data Engineering Bulletin, 2014.

- Identity Management for Converged Networks, a Sun/Lucent white paper, 2006.

- Sub-document queries over XML with XSQirrel, with Bogdan Alexe, WWW 2005.

- User Profile Management in Converged Networks (Episode II): "Share your Data, Keep your Secrets", with Bogdan Alexe, Irini Fundulaki, Pierre-Yves Lalilgand, Abdullatif Shikfa, Antoine Arnail, CIDR 2005.

- An Electronic Patient Record "on Steroids": Distributed, Peer-to-Peer, Secure and Privacy-conscious, with Serge Abiteboul, Bogdan Alexe, Omar Benjelloun, Bogdan Cautis, Irini Fundulaki, Tova Milo, VLDB 2004.

- ubQL: A distributed query language to program distributed query systems, PhD dissertation, 2001.

- Building intelligent Web applications using lightweight wrappers, with Fabien Azavant, in Data Knowledge Engineering, 2001; 354 citations.

- Everything You Ever Wanted to Know About DTDs, But Were Afraid to Ask (Extended Abstract), WebDB 2000.

EXTERNAL ENGAGEMENTS ↑

- **co-organizer** and **program chair** for Bloomberg Data for Good Exchange conference (2015–2019).

- *The future of High Performance Computing*, Feb 2014, **invited speaker** for Department of Energy.

- NYU Abu Dhabi International Hackathon for Social Good in the Arab world, Feb 2013, judge; Feb 2014, **co-organizer and judge**.

- *De Innovatione Rerum*, July 2013, Clermont-Ferrand, France, **invited speaker** for Michelin.

- Financial Empowerment Hackathon, March 2013, New York City, **judge**.

- NYC BigApps, March 2013, **invited speaker**.

- Reinvent NYC.gov, New York City government's first hackathon, July 2011, **winner** for user interface category.

- *Identity Management on Converged Networks: A Reality Check* panel at the WWW 2006 conference, **organizer**.

- academic research related duties: ACM SIGMOD 2009 (program committee), CIDR 2007 (chair and program committee).

PATENTS (12) ↑

- Managing information about entities using observations [US US9940381B1]

- Techniques for authenticating access to a private account at a public computing device using a user's mobile computing device [EP US EP2672417A2]

- Systems, methods, and computer-readable media for providing query suggestions [US US20160041991A1]

- Assessing ratings using tip patterns [US US20150170227A1]

- Proactive user-based content correction and enrichment for geo data [NL2009641C]

- Resource conscious tethering [US US8949439B1]

- Reducing power consumption of mobile device through automated environment [US US8504066B1]

- Managing information about entities using clusters of received observations [US US8706732B1]

- Machine-readable representation of geographic information [EP2292021A2]

- Ratings using machine-readable representation [EP2274687A4]

- Advertising [US US20180300745A1]

- Systematic approach to query optimization [US US6567802B1]